TECHNICAL RESEARCH REPORT

Representation of Spectral Profiles in the Auditory System
I: A Ripple Analysis Model

by S. Vranic - Sowers and S.A. Shamma

T.R. 94-14

Sponsored by
the National Science Foundation
Engineering Research Center Program,
the University of Maryland,
Harvard University,
and Industry

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
1994		**00-00-1994 to 00-00-1994**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Representation of spectral profiles in the auditory system. I: A ripple analysis model.	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Department of Electrical Engineering,Institute for Systems Research,University of Maryland,College Park,MD,20742**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**27**	

Representation of spectral profiles in the auditory system. I: A ripple analysis model.

Svetlana Vranić - Sowers
Department of Electrical Engineering and Institute for Systems Research,
University of Maryland, College Park, MD 20742, U.S.A.

Shihab A. Shamma
Department of Electrical Engineering, Institute for Systems Research,
and University of Maryland Institute for Advanced Computer Studies,
University of Maryland, College Park, MD 20742, U.S.A.

Received:

Short title: **A ripple analysis model.**

Abstract

A model of profile analysis is proposed in which a spectral profile is assumed to be represented by a weighted sum of sinusoidally modulated spectra (ripples). The analysis is performed by a bank of bandpass filters, each tuned to a particular ripple frequency and ripple phase. The parameters of the model are estimated using data from ripple detection experiments in [*Green*, 1986; *Hillier*, 1991]. Detection thresholds are computed from the filter outputs and compared with perceptual thresholds, for profile detection experiments with step, single component increment, and the alternating profiles. The model accounts well for the measured thresholds in these experiments. Physiological and psychophysical evidences from the auditory and visual systems in support of this type of a model are also reviewed. The implications of this model for pitch and timbre perception are briefly discussed.

INTRODUCTION

In characterizing the perception of spectral profiles, a basic objective is to select a model representation upon which an appropriate metric can be defined. Several such models have been proposed to account for data from a wide range of psychoacoustical tests – including profile analysis experiments and discriminations of simultaneous vowels. Examples are the independent channels model [*Durlach, Braida and Ito*, 1986], the maximum difference model [*Bernstein and Green*, 1987], the Ewaif model [*Feth, O'Malley and Ramsey*, 1982], the weighted slope model [*Klatt*, 1982], and the spectral peak model [*Assmann and Summerfield*, 1989]. Despite their unique characteristics, all models share the same fundamental starting point that the spectral profile is represented by the acoustic spectrum on a logarithmic frequency axis. Relative to this profile, various operations are defined to predict the perceptual thresholds.

In this paper, an alternative model of profile analysis is proposed which does not operate upon the profile directly, but rather upon its Fourier transformation. Specifically, it is hypothesized that an arbitrary profile is represented in the auditory system by a collection of weighted elementary sinusoidal profiles (ripples) of different frequencies and phases. Such a ripple analysis would be effectively accomplished via a bank of filters tuned to different ripple frequencies and phases. Detection thresholds would then be computed from this representation of the profile.

The primary motivations for such a ripple analysis model were findings from physiological mappings in the primary auditory cortex, AI [*Schreiner and Mendelson*, 1990; *Shamma et al.*, 1993]. The results revealed that AI potentially encodes, at each point along the tonotopic axis, an explicit measure of the local bandwidth and asymmetry of the acoustic spectrum. A broader interpretation of these two response measures (as we shall elaborate below) led to the notion that they may, respectively, correspond to the magnitude and phase of a Fourier transformation

of the profile. Recent neurophysiological results in AI are consistent with this hypothesis in that cortical cells are tuned to specific (characteristic) ripple frequencies and phases [*Calhoun and Schreiner*, 1993; *Shamma, Versnel and Kowalski*, 1993]. Furthermore, these two response properties are columnarly organized and topographically mapped along the isofrequency planes in a manner similar to that of the response area bandwidths and asymmetry eluded to above.

The idea that the brain analyzes and perceives its sensory patterns in this manner is relatively common in the vision literature where it has been variously called multi-resolution or multi-scale representation, and spatial frequency analysis [*Campbell and Robson*, 1968; *Levine*, 1985]. It is, however, the elegant anatomical and physiological work of [*DeValois and DeValois*, 1990] that has provided the most immediate inspiration to pursue this type of model for the auditory system.

In the psychoacoustical literature, it is a curious fact that the first doubts about the competence of the independent channels model were raised for ripple profile stimuli [*Green*, 1986]. Specifically, the detection thresholds were found to be remarkably constant and relatively high compared to what would be expected from the channel model. Furthermore, the thresholds could well be accounted for by the maximum difference model [*Bernstein and Green*, 1987] which assumes only two independent channels of information.

These issues will be examined here in the context of an auditory ripple analysis model. Essential to the development of such a model are basic psychoacoustical sensitivity measurements with simple ripple profiles. A few such experiments have already been performed [*Green*, 1986; *Hillier*, 1991; *Houtgast and Veen*, 1982]. Using these data and the conceptual framework outlined above, an explicit computational model is developed to interpret the results of various profile analysis experiments.

In the following sections, we first outline the computational model (Sec. I). Its various parameters are estimated in Secs. I C and II from experimental results reported here and in [*Green*, 1986; *Hillier*, 1991]. The model is used in Sec. III to predict the detection thresholds for several profile analysis experiments. Finally, the model is discussed in contrast to other profile analysis prediction models and in relation to auditory percepts such as timbre and pitch.

A. Terminology and notation

The following terms are frequently used here to describe the ripple analysis representation of profiles:

Ripple: refers to a sinusoidal spectral profile (e.g., $p(\omega) = \sin(2\pi\Omega\omega + \theta)$) on the logarithmic frequency axis ω. A ripple has a ripple frequency Ω (in cycle/octave) and a ripple phase θ (in radians or degrees).

Ripple spectrum, $P(\Omega)$: refers strictly to the Fourier transform of the profile $p(\omega)$.

Ripple analysis filter, $H(\Omega; \Omega_o, \Phi_o)$: refers to a bandpass filter that is tuned

around a characteristic ripple frequency (Ω_o) and phase (Φ_o).

Ripple transform, $r(\cdot)$: refers to the output of a bank of ripple analysis filters.

I. GENERAL DESCRIPTION OF THE RIPPLE ANALYSIS MODEL

The ripple analysis model can be conceptually divided into two stages. The first (Sec. I A) is a ripple transformation stage which simply converts the input spectral profile $p(\omega)$ into its corresponding ripple transform $r(\cdot)$. The second stage (Sec. II) is a detection model which operates on the magnitude or phase of the ripple transform, or on selected features of it such as its maxima and edges.

A. Computing the ripple transform of a spectral profile

This stage consists of a bank of ripple selective filters analogous to the frequency selective filters of the cochlea (Figs. 1). The impulse response of each filter, $h(\omega - \omega_o; \Omega_o, \Phi_o)$ (Fig. 1(a)), is centered around ω_o and is assumed to be a Gaussian shaped ripple of a particular ripple frequency Ω_o and ripple phase Φ_o ([*Gabor*, 1946]). In particular, the filters centered around $\omega_o = 0$ are defined as:

$$h(\omega; \Omega_o, \Phi_o) = 2\ g_o(\omega) \cos(2\pi\Omega_o\omega - \Phi_o), \tag{1}$$

where $g_o(\omega) = \sqrt{2\pi}\sigma\ e^{-\frac{(2\pi\omega\sigma)^2}{2}}$, and σ determines the width of the Gaussian envelope. In the Fourier transform domain, these filters are Gaussian shaped (Fig. 1(b) and (c)) and:

$$\begin{aligned} H(\Omega; \Omega_o, \Phi_o) &= H(\Omega; \Omega_o)\cos(\Phi_o) + \hat{H}(\Omega; \Omega_o)\sin(\Phi_o) \\ &= H(\Omega; \Omega_o)\ e^{-\jmath\ sign(\Omega)\Phi_o}, \end{aligned}$$

where $\hat{H}(\Omega; \Omega_o)$ is the Hilbert transform of $H(\Omega; \Omega_o)$ and $H(\Omega; \Omega_o)$ is the Fourier transform of $2\ g_o(\omega)\cos(2\pi\Omega_o\omega)$. Note that $H(\Omega; \Omega_o)$ ($\hat{H}(\Omega; \Omega_o)$) is pure real (imaginary).

In general, for each Ω_o and Φ_o there is a whole range of (identical) filters which are centered at different ω_o's along the tonotopic ω axis. Therefore, each impulse response is characterized by three parameters: Ω_o, Φ_o, and ω_o, and is given by: $h(\omega; \Omega_o, \Phi_o, \omega_o) = h(\omega - \omega_o; \Omega_o, \Phi_o)$.

The ripple transform at some ω_o is evaluated from the convolution of the impulse response and the input profile. For an arbitrary (real) input profile $p(\omega)$ with ripple spectrum $P(\Omega) = |P(\Omega)|e^{\jmath\theta(\Omega)}$, the response is:

$$\begin{aligned} r(\Omega_o, \Phi_o, \omega_o) &= \int_{-\infty}^{\infty} h(\omega_o - \omega; \Omega_o, \Phi_o)p(\omega)d\omega \qquad (2) \\ &= \int_{-\infty}^{\infty} H(\Omega; \Omega_o, \Phi_o)P(\Omega)e^{+\jmath 2\pi\Omega\omega_o}d\Omega \\ &= 2\int_0^{\infty} H(\Omega; \Omega_o)\ |P(\Omega)| \cdot \cos(\Phi_o - \theta(\Omega) - 2\pi\Omega\omega_o)d\Omega. \end{aligned}$$

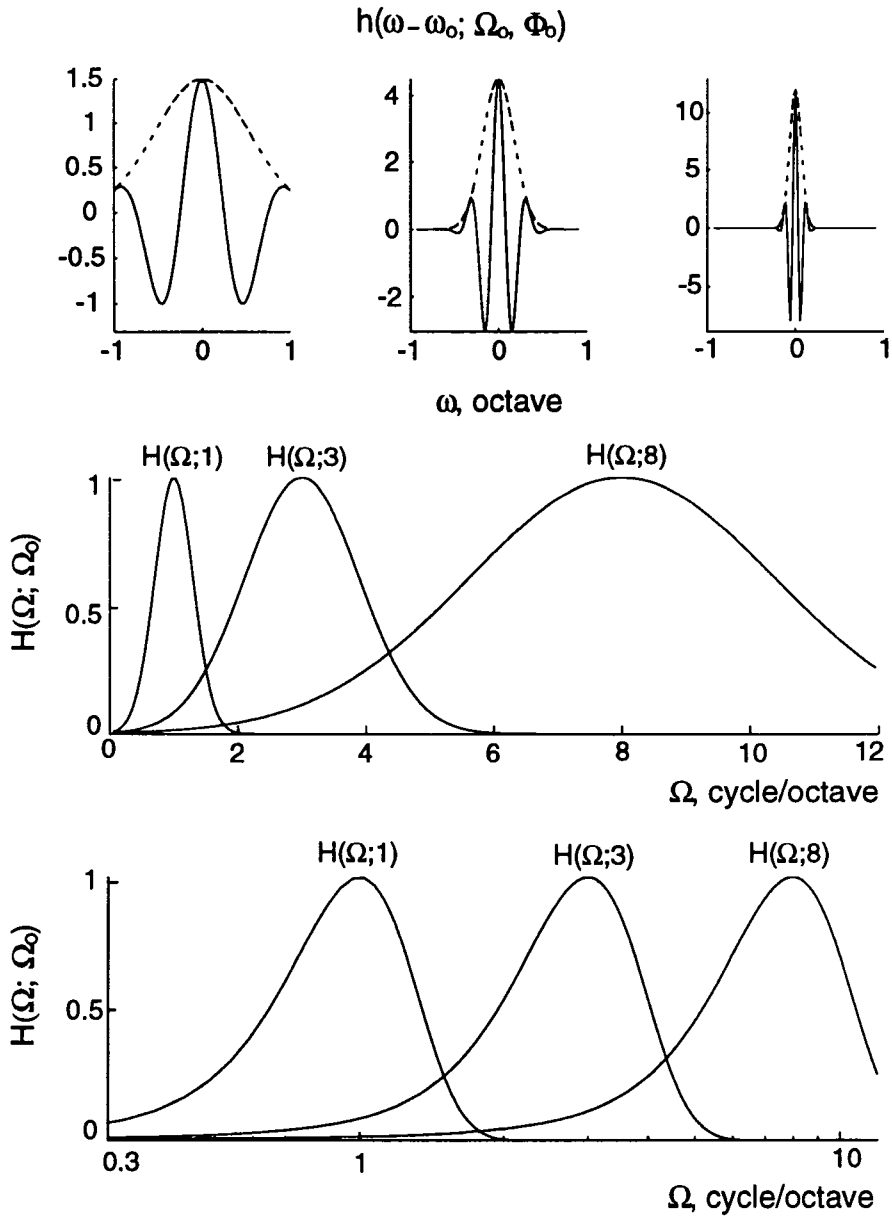

Figure 1: (a) Impulse responses of three filters with characteristic ripple frequencies $\Omega_o = 1, 3$, and 8 cycle/octave, and characteristic phase $\Phi_o = 0$. Filters are centered at $\omega_o = 0$ octave. The impulse response is computed as $h(\omega; \Omega_o) = 2\sqrt{2\pi}\sigma(\Omega_o)e^{-\frac{(2\pi\omega\sigma(\Omega_o))^2}{2}}\cos(2\pi\Omega_o\omega)$ for $\sigma(\Omega_o) = 0.3\ \Omega_o$. (b) Fourier transform of the three impulse responses of the filters in (a) plotted on a linear Ω axis. (c) Same as (b) but plotted on a logarithmic Ω axis.

This equation can be simplified for profiles which are even or odd symmetric around their center (arbitrarily designated at ω_o), or otherwise have a constant phase θ_o around it so that $\theta(\Omega) = sign(\Omega)\theta_o - 2\pi\Omega\omega_o$. This assumption applies approximately to all profiles discussed here and in the companion paper [*Vranić-Sowers and Shamma*, 19xx]. The ripple transform computed at such a point ω_o is:

$$r(\Omega_o, \Phi_o) = \cos(\Phi_o - \theta_o) \int_{-\infty}^{\infty} H(\Omega; \Omega_o)\ |P(\Omega)| d\Omega = \cos(\Phi_o - \theta_o)\ r(\Omega_o), \qquad (3)$$

where $r(\Omega_o)$ is a "magnitude" part:

$$r(\Omega_o) = \int_{-\infty}^{\infty} H(\Omega; \Omega_o)\ |P(\Omega)| d\Omega, \qquad (4)$$

and $\Phi_o - \theta_o$ is a "phase" part of the ripple transform. The majority of profile analysis tasks reported in the literature and considered here can be effectively described either as a change in the magnitude or in the phase of the ripple transform, as shown in Sec. II.

B. The representation of the input spectral profile

It is uncertain whether the auditory system represents its input spectral profile on a linear or a logarithmic amplitude scale [*Hillier*, 1991; *Shannon*, 1992]. For the model, the input profile $p(\omega)$ is taken to be the linear amplitude spectrum normalized by the amplitude of the base. Like the logarithmic spectrum, this representation is scale-normalized preserving only the level-independent features of the spectrum. Other possible inputs range from a simple scale-normalized power spectrum to more complicated biologically and psychoacoustically inspired representations, such as the excitation pattern model [*Glasberg and Moore*, 1990] and the auditory filter models of [*Hillier*, 1991; *Patterson*, 1986; *Shamma et al.*, 1986; *Slaney and Lyon*, 1990; *Yang, Wang and Shamma*, 1992].

In general, an inappropriate profile representation distorts its intended ripple spectrum $P(\Omega)$. In some cases the distortions are small, as demonstrated in Fig. 1(d) for spectral peaks or as in the case of small amplitude ripples where linear and logarithmic spectra look very similar and their perceptual thresholds are closely matched [*Green*, 1986; *Hillier*, 1991; *Houtgast and Veen*, 1982]. In other cases, the distortion is large but inconsequential in the context of the ripple analysis model as will be discussed in more detail in Sec. IV A.1.

C. Parameters of the filter bank

The filter bank depicted in Figs. 1(b) and (c) is assumed to be a constant-Q bank, i.e., its filters have constant widths on a logarithmic Ω axis or, equivalently, have linearly increasing widths (σ's) with Ω_o. This choice is primarily justified by data from adaptation experiments (both with ripples [*Hillier*, 1991] and visual gratings [*DeValois and DeValois*, 1990]) and neurophysiological experiments [*DeValois,*

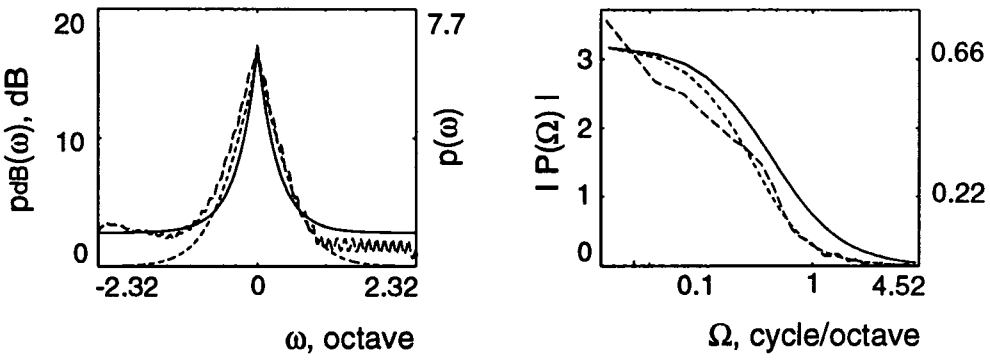

Figure 1: (d) Three input representations of a symmetric peak profile (left) and the corresponding ripple spectra magnitudes (right). There are little differences between the three representations or their ripple spectra. The solid line is the normalized (by the base) linear representation of the peak (right ordinate). The dotted line is the same peak profile represented on a logarithmic amplitude scale (left ordinate). The dashed line depicts the output of the excitation pattern model (no corrections were applied in the model, and the base was 0 dB amplitude; see [*Glasberg and Moore*, 1990] for details).

Albrecht and Thorell, 1982; *Shamma, Versnel and Kowalski*, 1993], in which filter tuning was estimated around various Ω_o's to be around 1 octave (measured at the half amplitude points). This corresponds approximately to $\sigma(\Omega_o) = 0.3\ \Omega_o$.

A fundamental consequence of the constant-Q property of the filters is that, a dilation (or a stretching) of the input profile (around ω_o) by a factor α, i.e., $p(\omega) \longrightarrow p(\alpha\ \omega)$ or $P(\Omega) \longrightarrow 1/\alpha P(\Omega/\alpha)$ causes only a simple translation of $r(\Omega_o)$ by $\log_2 \alpha$ octaves along the logarithmic Ω_o axis, leaving the shape of the ripple transform unaltered[1]. This property is illustrated in Figs. 2 for the ripple input at $\Omega_1 = 2$ cycle/octave and its dilated version ($\alpha = 1.5$) at $\Omega_2 = 3$ cycle/octave. The corresponding magnitudes of the ripple transforms are identical apart from a 0.58 octave shift (Fig. 2(c)). Note also that, a pure input dilation leaves the response as a function of Φ_o unaltered, i.e., it is largest at $\Phi_o = \theta_o$ as before.

[1]To see this, consider the effect on the ripple transform $r(\cdot)$ of dilating its input by a factor α. The response becomes:

$$r(\Omega_o, \Phi_o) = \cos(\Phi_o - \theta_o) \int_{-\infty}^{\infty} H(\Omega; \Omega_o)|P(\Omega/\alpha)|d\Omega/\alpha.$$

Evaluating $r(\cdot)$ at $\Omega_o/\alpha = \Omega_o'$ and letting $\Omega/\alpha = \Omega'$, we get:

$$r(\Omega_o', \Phi_o) = \cos(\Phi_o - \theta_o) \int_{-\infty}^{\infty} H(\Omega'; \Omega_o')|P(\Omega')|d\Omega',$$

which is identical in shape to $r(\Omega_o, \Phi_o)$ prior to dilation, except for a translation to Ω_o'.

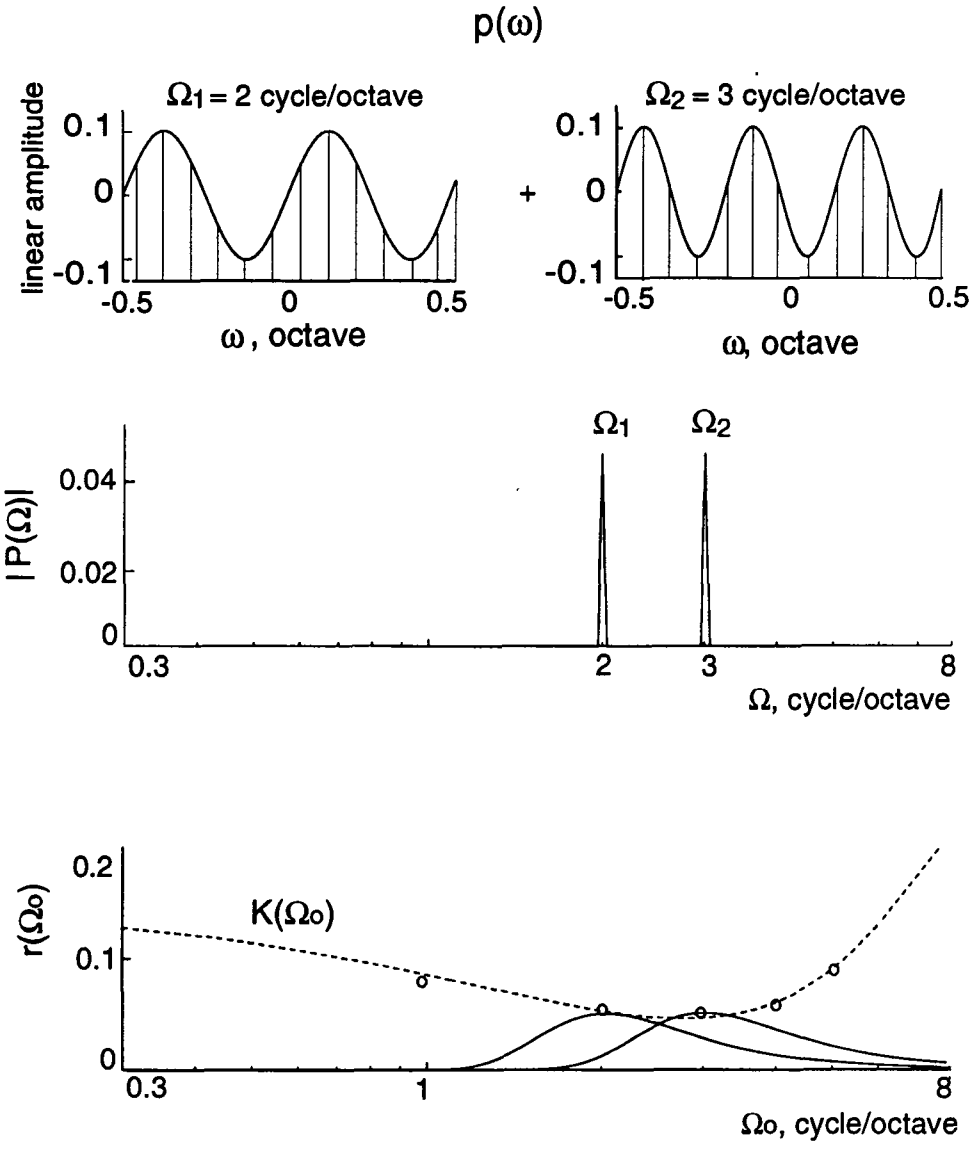

Figure 2: (a) Two ripple profiles with amplitudes 0.1 and frequencies $\Omega_1 = 2$ cycle/octave and $\Omega_2 = 3$ cycle/octave. (b) Ripple spectra magnitudes corresponding to the two ripple profiles. (c) Ripple transform magnitudes $r(\Omega_o)$ of the profiles in (a) (solid lines). The dashed line is a polynomial approximation to the measured data points $\mathbf{idl}(\Omega_o = \Omega)$ (denoted by circles) reproduced from Fig. 3.27 in [*Hillier*, 1991]. The detection threshold $K(\Omega_o)$ reflects the shape of the perceptual threshold $\mathbf{idl}(\Omega_o = \Omega)$. Maxima of $r(\Omega_o)$ are at their just detectable levels.

II. DETECTION PROCEDURES

Various detection procedures are developed in this section so as to predict perceptual thresholds from the ripple transform $r(\cdot)$. Two types of profile analysis paradigms will be distinguished: (**A**) Those in which the profile is to be detected against a flat standard, i.e., the task is to detect the existence of the profile. Most profile detection experiments fall in this category including those described by [*Bernstein and Green*, 1987; *Bernstein, Richards and Green*, 1987] and by [*Green*, 1986; *Hillier*, 1991]. The latter are the ripple detection experiments called here ripple intensity-difference-limen experiments or ripple-**idl**); (**B**) Those for which the standard is not flat. Instead, the subject is to detect a change in some parameter of an audible profile, for instance, in the frequency [*Hillier*, 1991] or phase of a ripple, or, in the height of a pedestal profile [*Green*, 1988].

A. Detection procedures for tests with flat standards

In such tests, the amplitude of the profile $p(\omega)$ is gradually increased until detection occurs. Therefore, in the context of the ripple analysis model, detection occurs when the magnitude of the ripple transform $r(\Omega_o)$ (according to Eq. (4)) exceeds some perceptual threshold level $K(\Omega_o)$ (Fig. 2(c)). In order to determine $K(\Omega_o)$, we use the results of the ripple-**idl** threshold measurements reported in [*Hillier*, 1991]. Figure 2(c) illustrates $r(\Omega_o)$ for two just detectable ripples Ω_1 and Ω_2 with amplitudes 0.1. In order to obtain the same detection results from the model, we define $K(\Omega_o) = \mathbf{idl}(\Omega_o = \Omega)$ as shown in Fig. 2(c)[2].

In Sec. III, we compute $r(\Omega_o)$ at perceptual thresholds for several profiles and compare them to $K(\Omega_o)$, in order to evaluate the performance of the model.

B. Detection procedures for tests with non–flat standards

Tests with non-flat standards involve threshold measurements of a parameter change in an audible profile. Detection procedures for three types of such tests are considered in this paper: (1) Dilation of a profile, exemplified by frequency-difference-limen (**fdl**) measurements for ripples profiles [*Hillier*, 1991]; (2) Ripple phase shift, such as the phase-difference-limen (**pdl**) measurements for ripple profiles (Sec. II B.2); (3) Change in overall amplitude, e.g., the pedestal-type experiments with single increment profiles [*Green*, 1988]. In relation to the ripple transform $r(\Omega_o, \Phi_o)$ in Eq. (3), these three tests correspond, respectively, to a translation in the magnitude of the ripple transform $r(\Omega_o)$, a translation in the phase $\Phi_o - \theta_o$, and a change in the amplitude of the ripple transform.

[2]This detection procedure is only the simplest of many possible schemes. For instance, one may assume K to be constant and instead weight the input profile or the filter heights by the inverse of the **idl** bowl [*Hillier*, 1991]. While these procedures are equivalent with respect to the single ripple **idl**'s, they generally have different consequences for arbitrary input profiles. In the absence of additional supporting data, we adopt the simplest approach taking K to be a function of Ω_o as shown in Fig. 2(c).

1. Detection of a profile dilation

As discussed earlier, dilation of a profile causes the magnitude of its ripple transform to translate along the logarithmic Ω_o axis without an overall change in shape. In the ripple analysis model, it is assumed that subjects detect this shift in $r(\Omega_o)$. Since the shift can be measured anywhere on the pattern, e.g., at its maximum or at its right or left edges, we choose to measure it in the steepest lowpass edge in $r(\Omega_o)$[3].

In order to predict the dilation thresholds in arbitrary profiles, we use the ripple-**fdl** measurements of [*Hillier*, 1991]. The data are reproduced by the dashed line in Fig. 3(c). The solid curve in Fig. 3(c) is the same data but translated by approximately 0.3 octaves (i.e., $\log_2 \frac{\Omega_1}{\Omega_1'}$, where $\Omega_1 = 0.8\Omega_1'$) to compensate for the misalingment between the location of the lowpass edge of $r(\Omega_o)$ relative to the ripple frequency Ω. The detectable shift is computed as $\Delta(\Omega_o) = \log_2(1 + \mathbf{fdl}(0.8\ \Omega_o))$ octave, where $\Omega = 0.8\Omega_o$ is the ripple frequency.

Therefore, the dilation threshold can be estimated for any arbitrary profile as follows: (1) Compute the magnitude of the ripple transform $|r(\Omega_o, \Phi_o)|$ of the profile; (2) Locate the steepest lowpass edge of the pattern along the Ω_o axis; (3) Determine the shift $\Delta(\Omega_o)$ based on the solid curve in Fig. 3(c) and the dilation threshold α from $\Delta = \log_2 \alpha$[4].

2. Detection procedure for phase shift in the ripple transform

In this test, the phase of the ripple transform, $\Phi_o - \theta_o$, is translated while holding the magnitude $r(\Omega_o)$ constant (see Eq. (3)). For an arbitrary profile, threshold is defined as the minimum detectable phase angle added to all components of the ripple spectrum. For a single ripple profile, this simply entails measuring the sensitivity to a phase shift in the profile, i.e., the phase-difference-limen of the ripple (**pdl**). Just as with the ripple-**fdl** measurements above, the ripple-**pdl** can be incorporated in the model to predict perceptual thresholds for arbitrary profiles. Experiments to obtain such **pdl** data are briefly described below.

[3]While other features of $r(\cdot)$ may be equivalent, our choice is motivated by the fact that the ripple transforms of arbitrary profiles are necessarily bandlimited (i.e., have lowpass edges) but may not always exhibit clear maxima.

[4]There are more explicit (and elaborate) schemes to incorporate the **fdl** curve into the model. For instance, one can set a constant detectable shift in the output and change the model parameters so as to require larger profile dilations to produce this (constant) output shift for ripples < 0.7 cycle/octave and > 6 cycle/octave (Fig. 3(c)). One way to accomplish this is to add a constant to the σ (e.g., $\sigma(\Omega_o) = 0.3\ \Omega_o + 0.05$). This increases the relative width of the filters substantially in the low Ω_o region (≤ 1 cycle/octave) in effect increasing the **fdl**'s in this range as observed in the data. Similarly, the **fdl** increase in the high Ω_o region (> 6 cycle/octave) may be related to the increasing **idl**'s there (Fig. 2(c)) and, hence, can be accounted for by introducing level-sensitive procedures for the detection of shifts in $r(\cdot)$.

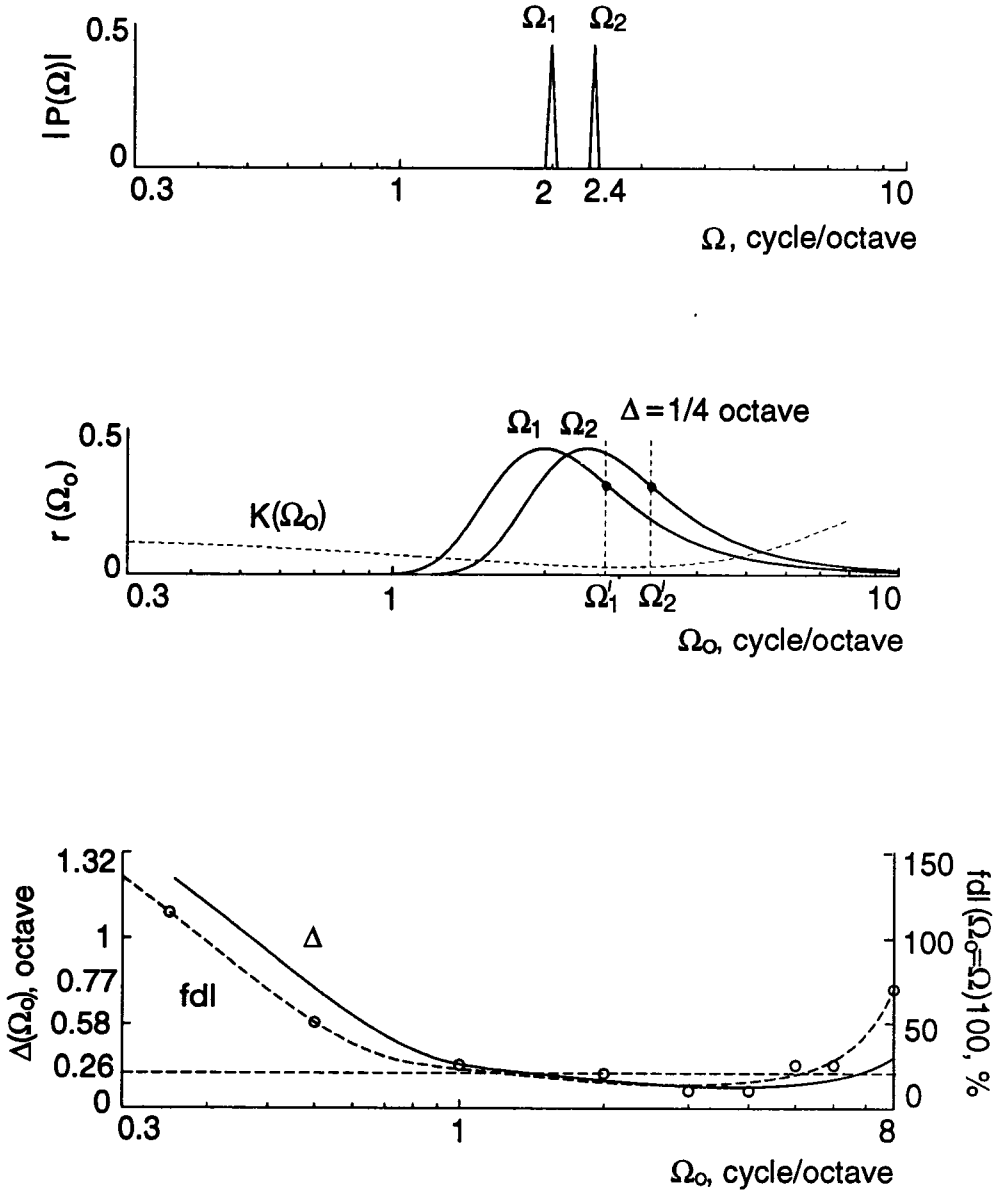

Figure 3: (a) Ripple spectra magnitudes of two ripple profiles at $\Omega_1 = 2$ cycle/octave and $\Omega_2 = 2.4$ cycle/octave. (b) Corresponding ripple transform magnitudes, $r(\Omega_o)$. Both ripples are well above their detectable levels (i.e., $r(\Omega_o)$ maxima significantly exceed $K(\Omega_o)$). The two ripple frequencies are 20% apart, which is the **fdl** threshold at $\Omega_1 = 2$ cycle/octave. This corresponds to a dilation factor of $\alpha = 1.2$ or a $r(\Omega_o)$ translation of $\Delta = \log_2 \alpha \approx 1/4$ octave. The dashed lines denote the locations of the steepest lowpass edges in $r(\Omega_o)$ (Ω'_1 and Ω'_2). (c) Dashed line is the interpolated **fdl**-threshold of data (denoted by circles) reproduced from Fig. 3.30 in [*Hillier*, 1991]. The values are shown on the right ordinate. Note that for the **fdl** curve, $\Omega_o = \Omega$ denotes the ripple frequency. Solid line represents the shifts (Δ) observed in the lowpass edges of $r(\Omega_o)$ corresponding to the same **fdl** data (see text for details).

(i) Methods

Sensitivity to ripple phase changes was measured in ripple profiles on a dB amplitude scale (Fig. 4(a)) and the **pdl** thresholds are reported in the units of degrees.

Sounds were generated at 25 kHz sampling rate via a Data Acquisition/Control Unit – HP3852A and two 16 bit 2-Channel Arbitrary Waveform DAC – HP44726A. They were low-pass filtered at 10 kHz and passed through an equalizer (IEQ One/Third Octave Intelligent Programmable) for level adjustment. Before presentation to listeners, sounds were gated for a 110 ms duration including 10 ms rise and decay ramps. Sounds were delivered inside an acoustic chamber through a speaker (ADS L470), i.e., without headphones.

A "two-alternative two-interval" forced choice adaptive procedure was used to estimate the thresholds. Each trial consisted of two 110 ms long observation intervals separated by 500 ms pause. After listener's response, a short visual feedback was provided and a new trial started until all 50 trials that comprise one block were presented.

The discrimination task was to distinguish between the *standard*, which did not change over a block of trials, and the *signal*, which resembled the *standard* except for an adaptive change in the ripple phase in each trial. The step size was defined in terms of a change in the ripple phase and it varied from $0.6^o - 4^o$ depending on the testing condition.

On the first trial the signal was three step sizes away from the standard. On each subsequent trial the signal was changed according to the "two down-one up" procedure in order to estimate the level that produces 70.7% correct answers ([*Levitt*, 1971]). The step size was halved after 3 reversals and the threshold was estimated as the average of the signal across the last even number of reversals excluding the first three. Signal and standard occurred with equal *a priori* probability in one of the two intervals.

The overall presentation level was randomized across trials and within a trial over a 20 dB range in 1 dB resolution, in order to ensure that listeners based their judgement on a change in spectral shape rather than on absolute level change in a particular frequency band ([*Green*, 1988]).

The results reported are based on data from two normal hearing subjects. Subjects were trained for about a week (four days a week, 60 – 90 minutes per day) before the actual recording took place.

(ii) Stimulus

For all testing conditions the number of frequency components was 161 (34 per octave) and the frequency components were equally spaced on a logarithmic scale between 0.2–5 kHz. The starting ripple phase was kept constant at zero degrees for the data reported here. Other starting phases were also tested and results were very similar. The ripple frequency (Ω) was fixed over a set of trials. The peak-to-

valley ratio was defined as $20 \log \frac{a_{max}}{a_{min}}$, where a_{max} and a_{min} are the peak and valley amplitudes of the profile (Fig. 4(a)).

(iii) Results

The average data for two subjects are presented in Fig. 4(b) as a function of ripple frequency for two ripple levels. The results show that **pdl**'s are constant below about 2 cycle/octave at both levels tested, achieving a minimum of about 6° for the larger level. Phase sensitivity decreases with increasing ripple frequencies beyond 2 cycle/octave.

Figure 4(c) depicts the data for individual subjects as a function of ripple level. Thresholds saturate with increasing level at all ripple frequencies tested.

(iv) Discussion

Data in Fig. 4(b) suggest that, at low Ω's, subjects detect a constant phase shift and not a constant displacement of the peaks, as is probably the case for higher than 2 cycle/octave ripple frequencies. Since the slope of the **pdl** curves for $\Omega > 2$ cycle/octave (Fig. 4(b)) is approximately 3.8° per octave, then the constant positional shift can be estimated to be approximately 0.73% (or 0.01 octaves).

In summary, given any arbitrary profile whose ripple transform contains significant energy below 1 cycle/octave, the model predicts a constant phase shift detection threshold of approximately 6°. Thresholds should slowly begin to increase if the ripple transform is shifted to ripple frequencies greater than 1 cycle/octave.

3. Detection procedure for pedestal-type experiments

In pedestal-type experiments, an audible profile is increased in amplitude until the change is detected. Since the ripple analysis model presented here is linear, then the ripple transform of the profile increases proportionately with the input profile. In order to specify the detection thresholds, it is necessary to collect **idl**-like data for different ripples at various ripple pedestals. Such data are not available at present.

As an approximate measure, one can utilize the data from single component pedestal experiments [*Green*, 1988]. This narrow profile has a broadband ripple transform and, hence, its detection threshold might be assumed to be due to the most sensitive ripple components. Thus, one can at least predict that arbitrary profiles with similarly broad ripple transforms should exhibit comparable thresholds. This conjecture will be tested later in Part II of this paper [*Vranić-Sowers and Shamma*, 19xx].

C. Stochastic detection procedures

The filter bank of the ripple analysis model can be viewed as a set of independent channels conveying information about the ripple spectrum of the profile $P(\Omega)$. In this sense, it is analogous to the classical view of the critical band channels operating upon the spectral profile. Hence, the independent channels model of [*Durlach, Braida and Ito*, 1986] and the more specific model of [*Bernstein and Green*, 1987]

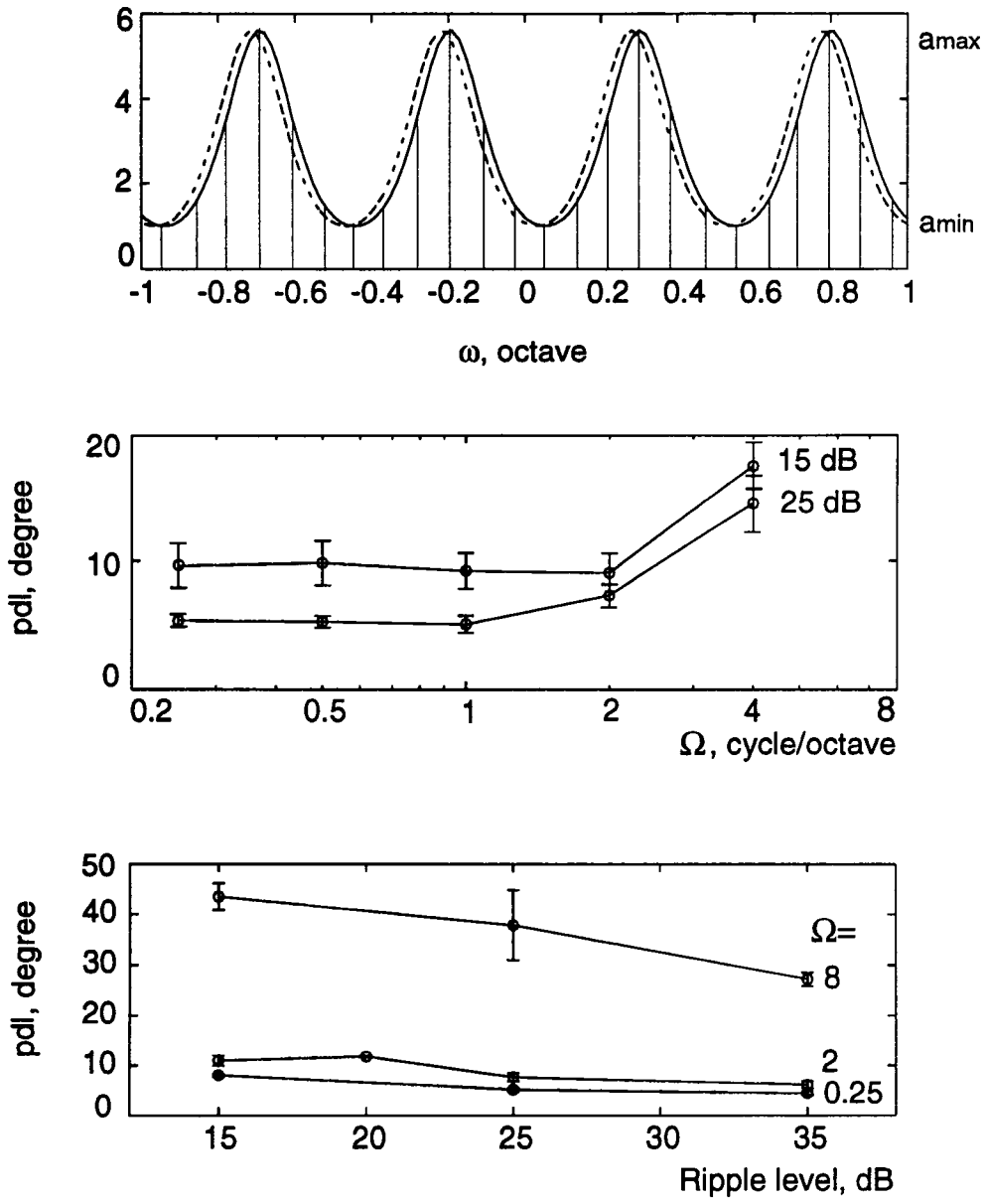

Figure 4: (a) A sinusoidal ripple profile with $\Omega = 2$ cycle/octave and 15 dB peak-to-valley amplitude (computed as $20\log_{10}(a_{max}/a_{min})$). Dashed line is its 16^o phase shifted version. (b) Phase difference limen threshold (**pdl**) as a function of ripple frequency for 15 dB and 25 dB peak-to-valley amplitudes (or ripple levels) averaged over 2 subjects. (c) Individual **pdl** thresholds at three ripple frequencies as a function of ripple level. (Subject 1 was tested at 0.25 and 8 cycle/octave, and subject 2 at 2 cycle/octave and at additional 20 dB ripple level).

can be formally adapted and applied to the outputs of the ripple filter bank. This strategy is not pursued here because of the lack of sufficient data on such parameters of the channels as their variances.

D. Summary of the ripple analysis model

The ripple analysis model consists of the following computational steps:

(1) Compute $P(\Omega)$, the ripple spectrum of the input profile $p(\omega)$ as: $P(\Omega) = \int p(\omega)e^{-j2\pi\Omega\omega}d\omega$.

(2) Compute $r(\Omega_o, \Phi_o)$, the output of the filter bank using Eq. (2) (or, in the special case Eq. (3)). The width of the filter $H(\Omega;\Omega_o)$ centered at Ω_o is determined from $\sigma(\Omega_o) = 0.3\ \Omega_o$.

(3) For *flat standard* profile experiments, compare the magnitude of the ripple transform $|r(\Omega_o, \Phi_o)|$ to the perceptual threshold curve $K(\Omega_o)$ as shown in Fig. 2(c).

(4) For *non–flat standard* profile experiments, three types of tests are considered in the model: (a) **fdl**-*type tests*: Threshold is estimated from $\Delta(\Omega_o)$ in Fig. 3(c), where Ω_o is the location of the steepest lowpass edge in $|r(\Omega_o, \Phi_o)|$; (b) **pdl**-*type tests*: Threshold is estimated from the **pdl** curve in Fig. 4(b); (c) *pedestal-type experiments*: For broad profiles, threshold is the smallest detectable percentage change in $|r(\Omega_o, \Phi_o)|$. It is assumed to be approximately equal to the single component pedestal thresholds given in Fig. 5.4 in [*Green*, 1988].

III. PREDICTIONS OF THE RIPPLE ANALYSIS MODEL FOR VARIOUS INPUT PROFILES

In this section, the model is used to account for the perceptual thresholds measured in several profile analysis experiments. Since the ripple profile data of [*Green*, 1986; *Hillier*, 1991] have been used to estimate the model parameters, the only other reported profile experiments that we can consider here are of the **idl**-type. These are the detection of three profiles against a flat standard: the alternating, the step, and the single component increment profiles. In Part II of this paper, we shall present new data to test model predictions in the three examples of non–flat standard tests outlined above.

A. Predicting detection thresholds for an alternating profile

The alternating profile [*Bernstein and Green*, 1987] consists of 21 uniformly distributed components (0.2–5 kHz) that alternate above and below a flat (unit) base (Figs. 5). Thresholds for detection of such a profile are reported at –21.7 dB ($= 20\log(\delta a) = 20\log 0.08$), where δa is the amplitude of an alternating component relative to the unit base. Such a profile can be considered approximately a ripple at the highest possible frequency representable by this complex, i.e., at 10 cycles per 4.64 octaves, or approximately 2.15 cycle/octave. The amplitude δa of the just detectable ripple at this frequency can be predicted from $K(\Omega_o)$ as $\delta a = 2K(\Omega_o =$

2.15 cycle/octave) $\approx$ 0.1 (or -20.0 dB), which is close to the measured amplitude (Fig. 5(c)).

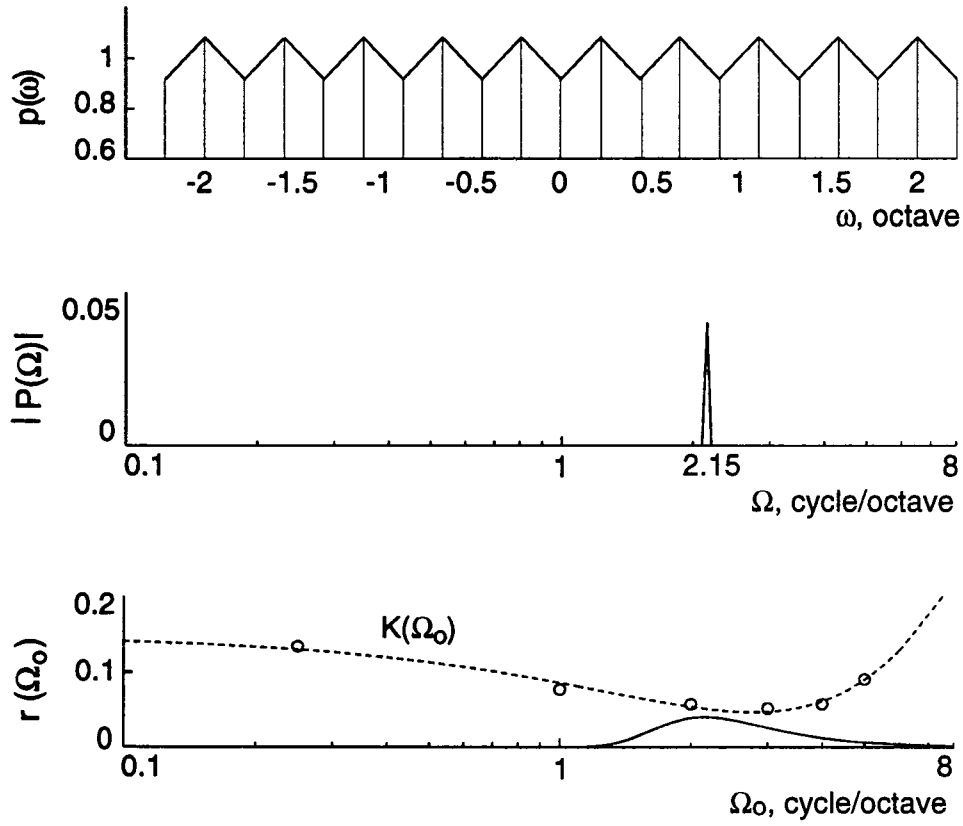

Figure 5: (a) The alternating profile at threshold amplitude (0.08, or -21.7 dB). (b) Ripple spectrum magnitude of the alternating profile in (a). (c) Ripple transform magnitude of the profile in (a). The detection threshold $K(\Omega_o)$ in (c) is reached near 2.2 cycle/octave.

B. Predicting detection thresholds for a step profile

The task in this experiment is to detect the presence of a (linear) step in a 21 component flat standard [*Bernstein and Green*, 1987] (Fig. 6(a)). For a step-up that is centralized (located at 1 kHz), threshold is reached at -23.1 dB ($= 20 \log \delta a = 20 \log 0.07$), where δa is the height of the step (relative to the unit base). Figure 6(b) (solid line) illustrates the ripple spectrum of this (idealized) profile. The corresponding model output $r(\Omega_o)$ is a constant because of the constant-Q property of the filters[5](Fig. 6(c), solid line). The predicted threshold is smaller than measured (0.05 or -26 dB). However, a more realistic representation of the step is with a gradual (or ramped) transition because of cochlear filter smoothing (dashed lines in Figs. 6). The smoothing of the ideal profile lowers the $P(\Omega)$ (Fig. 6(b))

and the corresponding $r(\Omega_o)$ is more lowpass filtered and just detectable near 2 cycle/octave. Note, however, that the model predicts that more heavily smoothed step profiles should become less detectable.

Since the phase of the ripple spectrum does not play a role here, predicted thresholds remain the same for the reversed (step-down) profile as is indeed measured. Finally, the simplified model cannot account for the rise in thresholds as the step is moved towards the edges of the spectrum [*Bernstein and Green*, 1987]. It may be possible, however, to account for this trend by including the effects of the base edge in $p(\omega)$ and by using the complete model (i.e., Eq. (2)).

C. Predicting detection thresholds for a single component increment profile

In this experiment, a single component in the profile is incremented relative to the base [*Green*, 1988] (Fig. 7(a)). The threshold is approximately -20.1 dB ($= 20\log\delta a = 20\log 0.09$), where δa is the height of component relative to the (unit) base. In order to apply this profile to the ripple analysis model, it is assumed that the finite bandwidth of the cochlear filters broadens the impulse-like profile, making it appear as a narrow symmetric peak profile, e.g., with approximately 0.1 octave bandwidth (measured at the 3 dB points) with same height as before (= 0.09). Figure 7(c) illustrates that for such a peak the corresponding output $r(\Omega_o)$ just reaches perceptual threshold $K(\Omega_o)$ near $\Omega_o = 2.3$ cycle/octave. Note that, approximating the increment by a peak with slightly different bandwidths causes small shifts in the broad $r(\Omega_o)$ without affecting the above estimated thresholds significantly.

IV. DISCUSSION

A. Summary of the ripple analysis model and underlying assumptions

A simplified ripple analysis model is presented to integrate findings from various profile analysis experiments. The basic operation implied by the model is a transformation of the profile into its ripple transform domain. Various manipulations on the profile are then interpreted and detected in this domain. Two sets of assumptions underlie the model: the nature and linearity of the input representation and the parameters of the ripple analysis filters.

[5]This assertion can be verified as follows. Consider a constant-Q filter ($\sigma(\Omega_o) = \sigma_{rel}\ \Omega_o$). For a step profile input (with magnitude of ripple spectrum $|P(\Omega)| = 1/|\Omega|$), the filter outputs $r(\cdot)$ are given by:

$$r(\cdot) = 2\int_{\epsilon}^{\infty} \frac{1}{\Omega} e^{-\frac{(\Omega-\Omega_o)^2}{2\sigma^2(\Omega_o)}} d\Omega = 2\int_{\epsilon}^{\infty} \frac{1}{\Omega} e^{-\frac{(1-\Omega/\Omega_o)^2}{2\sigma_{rel}^2}} d\Omega = 2\int_{\epsilon}^{\infty} \frac{1}{\Omega'} e^{-\frac{(1-\Omega')^2}{2\sigma_{rel}^2}} d\Omega',$$

where $\Omega' = \Omega/\Omega_o$ and ϵ is a small positive number. Therefore, $r(\cdot)$ is a function of σ_{rel} only and, specifically, it is independent of Ω_o.

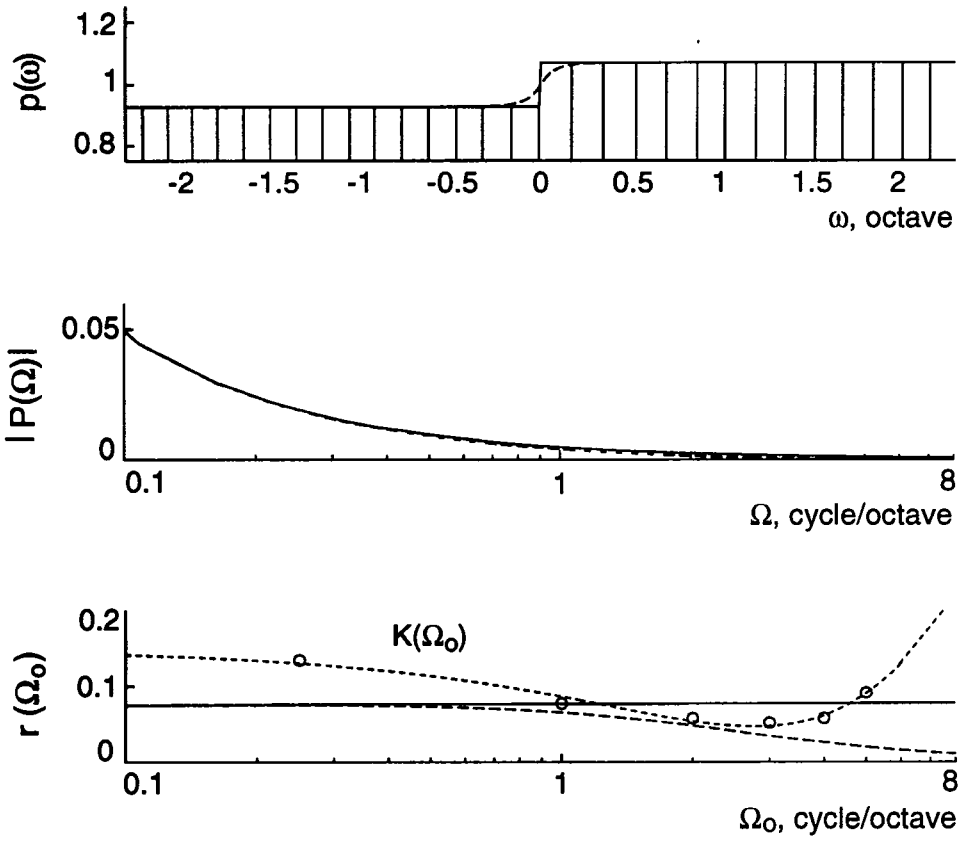

Figure 6: (a) Profile of a step function (solid line) at threshold amplitude (0.07, or –23.1 dB) and its smoothed version (dashed line). (The smoothed version is obtained by convolving the step with the narrow symmetric peak profile of bandwidth 0.1 octave and amplitude –30 dB with respect to the base). Magnitudes of (b) ripple spectra and (c) ripple transforms corresponding to the idealized (solid lines) and smoothed (dashed lines) step. The ripple spectrum magnitude of the idealized step is computed as: $|P(\Omega)| = 2 \cdot \delta a / |j\ 2\pi\Omega\ M|$, where δa is the amplitude at threshold.

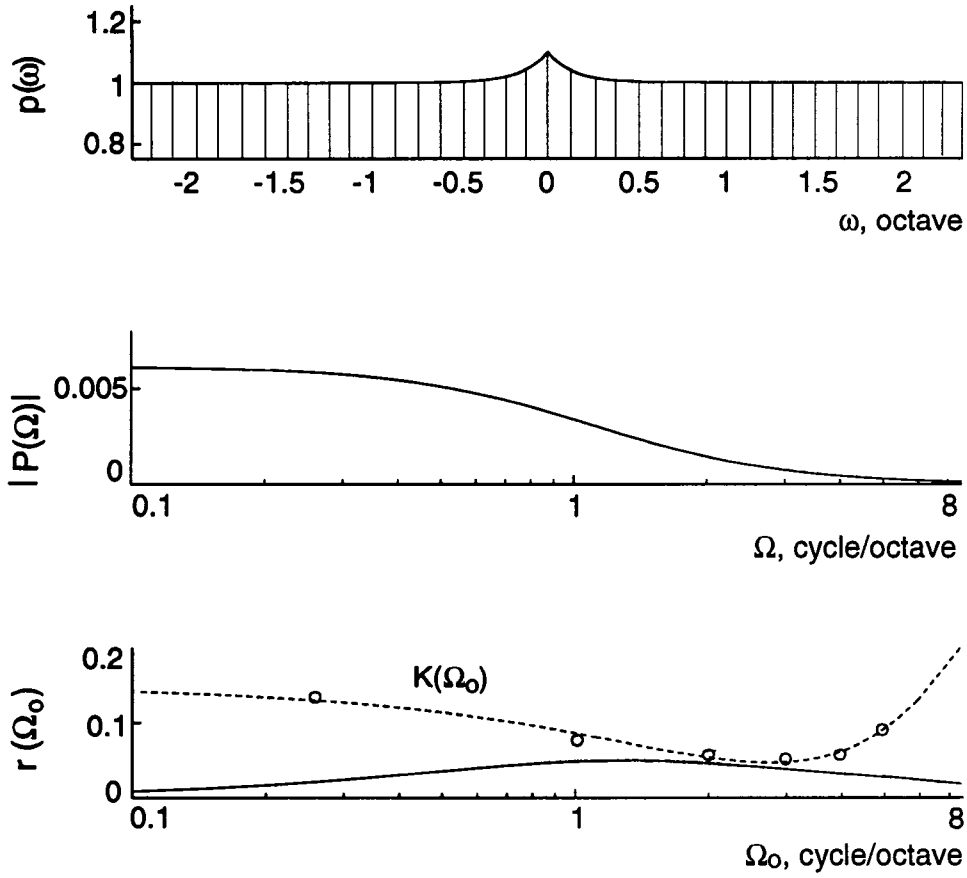

Figure 7: (a) A smoothed version of a single increment profile on a flat base. The amplitude is set at its perceptual threshold (0.09, or –20.1 dB). The single increment is approximated with the narrow symmetric peak of 0.1 octave bandwidth and –20.1 dB amplitude. The ripple spectrum and ripple transform magnitudes are shown in (b) and (c), respectively.

1. The representation and linearity of the input profiles

It is assumed in this model that the auditory system analyzes the amplitude spectrum on a linear, rather than on a logarithmic, scale. Neither is known to be the true auditory representation and other representations such as the power spectrum or some output of a cochlear filter model might be more appropriate. The effects of using a distorted representation are minimal in the cases examined in this paper because it usually creates distortion components of smaller amplitudes that, for **idl** tests, are effectively sub-threshold at their corresponding filters.

The exact nature of the input profile is more consequential in cases where metrics between different complex profiles are considered (see discussion later) or when profiles are added linearly. This brings up a fundamental assumption of the ripple analysis model, that the auditory system analyzes linearly a profile in terms of ripples. How does the cochlea with all of its nonlinearities preserve the principle of superposition of spectral ripples? And if not, in what form is the linearity preserved so as to permit this type of ripple analysis? Hillier attempted to address this issue using adaptation experiments [*Hillier*, 1991]. Recent models of cochlear processing have also tackled this question [*Wang and Shamma*, 1994]. However, the validation of the linearity hypothesis must await direct tests from psychoacoustical and neurophysiological experiments in search of systematic interactions among a small number of simultaneously presented ripples.

2. The parameters of the ripple analysis filters

The filters determine the shape of the ripple transform $r(\Omega_o, \Phi_o)$ and, hence, the interpretation of the results. The choices made here regarding the parameters and shape of these filters satisfy two basic experimental findings reported in [*Hillier*, 1991] (Secs. 4.4 and 4.5) and in psychophysical experiments in vision using analogous stimuli (summarized in [*DeValois and DeValois*, 1990]). These are: (1) the filters are roughly of a constant-Q factor, and (2) their width is approximately 1 octave (i.e., $\sigma(\Omega_o) = 0.3\ \Omega_o$). It might be argued that other details of the filter shapes (such as their heights and form) can be inferred from the **idl** and **fdl** measurements. Such an inference, however, is uncertain as discussed in Sec. II since other parameters can be readily adjusted with similar effects on the model outputs. To avoid making such specific commitments in the model, the filters are defined in as general and simple a form as possible.

B. Distinguishing the ripple analysis model from other profile analysis models

The computations outlined in this paper served to illustrate the competence of the ripple analysis model in accounting for several profile analysis measurements. However, other models such as the independent channels models [*Durlach, Braida and Ito*, 1986] and the maximum difference model [*Bernstein and Green*, 1987] can account for a significant portion of the same data. It is, therefore, important to

come up with specific tests that can distinguish these models. Two such tests follow from the fundamental predictions which emerge from the ripple analysis model:

Consider any arbitrary spectral profile whose ripple transform magnitude $|r(\Omega_o, \Phi_o)|$ *is large relative to the perceptual threshold* $K(\Omega_o)$. *Then:*

(1) If $|r(\cdot)|$ *has a lowpass edge approximately in the range 0.85–7 cycle/octave, dilation thresholds are predicted to be constant at 20–30%.*

(2) If $|r(\cdot)|$ *contains at least some large low frequency ripples (*< 1 *cycle/octave), phase sensitivity is expected to be constant at approximately* 6^o.

Both of these predictions are unintuitive and, hence, their confirmation reflects well on the model. They are directly tested in the companion paper [*Vranić-Sowers and Shamma*, 19xx], where dilation and phase thresholds are measured for a complex (i.e., not a single ripple) profile.

C. Relevance to timbre perception

So far we have focused on the model output around the center of a profile at ω_o. In general, for an arbitrary profile $p(\omega)$ the simplifications leading to Eq. (4) do not necessarily apply. In that case, a spectral profile should be expanded along all three independent axes: Ω_o, Φ_o, and ω_o.

An important aspect of the complete representation is its locality with respect to the tonotopic axis. This is analogous to the locality in time of a spectrogram of running speech. Computationally, the locality of the ripple analysis is implied by the relatively broad bandwidths of the ripple filters or, equivalently, the limited extent of the impulse response of the filters (Fig. 1(a)).

Since changes along any of the three axes are perceptually detectable, then comparing two arbitrary profiles must be done along all three dimensions of the ripple representation. Metrics based on this representation (e.g., simple Eucledian distance) might be considerably simpler than other metrics (based on the spectral domain representation), since other metrics often imply "conditioning" operations which are included in the ripple transform representation. For example, the metric suggested by [*Assmann and Summerfield*, 1989] applies (among other things) a second derivative upon the profile, effectively de-emphasizing the slow variations (or, equivalently, the low frequency ripples) of the profile. Such an operation is implied in the model by the "highpass" (left) edge of the **idl** curve.

D. Relation to rippled noise stimuli and the pitch of complex sounds

A different rippled spectrum stimulus that has been widely used in studies of pitch perception is the so-called *rippled noise* [*Yost and Hill*, 1979]. It has a sinusoidal spectral envelope defined against a linear, rather than a logarithmic frequency, axis, i.e., is similar to a harmonic spectrum. On a logarithmic frequency axis, however, a harmonic spectrum appears to have an exponentially increasing ripple frequency.

The representation of harmonic spectra in the ripple analysis model leads to

many interesting hypotheses regarding the encoding of complex pitch in the auditory system. Of immediate relevance to the focus of this paper, however, is the interpretation of the "dominance region" in pitch models. Specifically, it has long been known that the 2^{nd}, 3^{rd}, and 4^{th} harmonics in a series are dominant in conveying the pitch of the complex [*Plomp*, 1976]. From a computational point of view, pitch models have accounted for this phenomenon by emphasizing (or weighting more heavily) these regions of the spectral profile prior to estimating the pitch strength and value [*Yost and Hill*, 1979].

The dominance region can be viewed in the context of the ripple analysis model as a correlate of the ripple-**idl** sensitivity curve (Fig. 2(c)), which has its lowest thresholds for ripples around 2 cycle/octave. In a harmonic spectrum defined against a logarithmic axis, the spectral profile around the $2^{nd} - 4^{th}$ harmonics has the same ripple frequencies. Thus, the emergence of the ripple-**idl** curve may share the same origins as those responsible for the dominance region, namely a combination of suppressive and other interactions at relatively peripheral stages of the auditory system [*Bilsen et al.*, 1975; *Wang and Shamma*, 1994; *Yost and Hill*, 1979].

E. Relation to visual processing

An appealing aspect of the ripple analysis model is that it shares the conceptual framework of *spatial frequency analysis* that has long been prevalent in visual processing. While having its critics (see reviews in [*DeValois and DeValois*, 1990]), this approach has been supported by substantial anatomical, neurophysiological, and psychophysical evidence, elegantly detailed in [*DeValois and DeValois*, 1990]. Interestingly, in the vision community, the idea that the brain performs a local Fourier transformation is motivated by its similarity to the cochlear transformations of the auditory system! From the perspective of the auditory system, however, the cochlea and early auditory stages simply transform a sound into an input spectral profile. The auditory nervous system then treats this profile the same way the visual system treats its retinal image.

The notion of a Fourier transformation of a spectrum is common in engineering speech applications and is known as convolutional homomorphic processing. It involves computing Fourier-like coefficients of the spectral profile known as cepstral coefficients [*Oppenheim and Schafer*, 1990]. While quite different in details, the cepstral coefficients encode roughly similar types of information about the shape of the spectrum as the ripple transform, and have been found especially useful in automatic speech recognition systems.

Finally, the correspondence between the auditory ripple analysis and the visual spatial frequency analysis goes deeper than a mere analogy. As evidence to this claim, consider the closely matched values of the filter parameters and detection thresholds measured in the auditory and visual systems, e.g., roughly constant-Q and 1 octave wide filters, and the constant 6^o phase sensitivity which increases at higher ripple frequencies (Table 6.1 and Figs. 6.11 and 8.3 in [*DeValois and DeVal-*

ois, 1990]). These remarkable equivalences may simply reflect modality-independent limitations imposed by identically structured sensory areas in the central nervous system. For instance, the resolution of the analysis filters may be dictated by developmental rules limiting the minimum divergence or convergence of dendritic fields along the sensory epithelium. Clearly, exploring further equivalences and differences between similarly defined psychophysical measures, e.g., **fdl**'s for ripples *versus* gratings (which apparently have not been reported in the literature), would shed considerable light on the underlying functional organization of both systems.

ACKNOWLEDGMENTS

This work was supported by grants from the Air Force Office of Scientific Research, The Office of Naval Research, and the National Science Foundation (NSFD CDR8803012). The authors are members of the Institute for Systems Research. We are grateful to Dr. David Green and two anonymous reviewers for extensive reviews and helpful suggestions on the manuscripts, and to Daniel Naar from the Apple Corporation for valuable discussions.

REFERENCES

Assmann, P. F. and Q. Summerfield, Modeling the perception of concurrent vowels: Vowels with the same fundamental frequency, *J. Acoust. Soc. Am.*, *85*, 327–338, 1989.

Bernstein, L. R. and D. M. Green, Detection of simple and complex changes of spectral shape, *J. Acoust. Soc. Am.*, *82(5)*, 1587–1592, 1987.

Bernstein, L. R., V. M. Richards and D. M. Green, in *Auditory Processing of Complex Sounds: The detection of spectral shape change*, pp. 6–15, Lawrence Erlbaum Associates, Inc., New Jersey, 1987.

Bilsen, F. A., J. H. ten Kate, T. J. F. Buunen and J. Raatgever, Response of single units in the cochlear nucleus of the cat to cosine noise, *J. Acoust. Soc. Am.*, *58*, 858–866, 1975.

Calhoun, B. M. and C. E. Schreiner, Spatial frequency filters in cat auditory cortex, *Neurosc. Meeting, November*, *19*, 581.8, 1993.

Campbell, F. W. and J. G. Robson, Application of Fourier Analysis to the visibility of gratings, *J. Physiol.*, *197*, 551–566, 1968.

DeValois, R. L., Albrecht and Thorell, *Vision Research* , *22*, 545–559, 1982.

DeValois, R. L. and K. K. DeValois, in *Spatial Vision*, Oxford University Press, New York, 1990.

Durlach, N. I., L. D. Braida and Y. Ito, Towards a model for discrimination of broadband signals, *J. Acoust. Soc. Am.*, *80 (1)*, 63–72, 1986.

Feth, L. L., H. O'Malley and J. J. Ramsey, Pitch of unresolved two-component complex tones, *J. Acoust. Soc. Am.*, *72*, 1403–1412, 1982.

Gabor, D., Theory of communications, *J. Inst. Elec. Eng. (London)*, *93*, 429–457, 1946.

Glasberg, B. R. and B. C. J. Moore, Derivation of auditory filter shapes from notched-noise data, *Hearing Res.*, *47*, 103–138, 1990.

Green, D. M., 'Frequency' and the detection of spectral shape change, in *Auditory Frequency Selectivity*, edited by B. C. J. Moore and R. D. Patterson, pp. 351–359, Plenum Press, Cambridge, 1986.

Green, D. M., in *Profile Analysis*, Oxford Press, New York, 1988.

Hillier, D. A., Auditory processing of sinusoidal spectral envelopes, Ph.D. Dissertation, The Washington University and Severn Institute, 1991.

Houtgast, T. and T. M. van Veen, On the just-detectable modulation of the spectral envelope on a log-f scale, *J. Acoust. Soc. Am. Suppl. 1*, *71*, S37, 1982.

Klatt, D. H., Prediction of perceived phonetic distance from critical-band spectra: A first step, *Proc. ICASSP*, *2*, 1278–1281, 1982.

Levine, M., *Vision in Man and Machine*, McGraw-Hill, New York, 1985.

Levitt, W., Transformed Up-Down Methods in Psychoacoustics, *J. Acoust. Soc. Am.*, *49*, 467–477, 1971.

Oppenheim, A. and R. Schafer, in *Digital Signal Processing*, Prentice-Hall, New Jersey, 1990.

Patterson, R. D., Auditory filters and excitation patterns as representations of frequency resolution, in *Auditory Frequency Selectivity*, edited by B. C. J. Moore and R. D. Patterson, pp. 123–177, Plenum Press, Cambridge, 1986.

Plomp, R., in *Aspects of Tone Sensation*, 1976.

Schreiner, C. and J. Mendelson, Functional topography of cat primary auditory cortex: Distribution of integrated excitation., *J. Neurophysiol.*, *64(5)*, 1442–1459, 1990.

Shamma, S. A., R. S. Chadwick, J. Wilbur, K. A. Moorish and J. Rinzel, A biophysical model of cochlear processing: Intensity dependence of pure tone responses, *J. Acoust. Soc. Am.*, *80(1)*, 133–145, 1986.

Shamma, S. A., J. W. Fleshman, P. R. Wiser and H. Versnel, Organization of Response Areas in Ferret Primary Auditory Cortex, *J. Neurophysiol.*, *69* , 367–383, 1993.

Shamma, S. A., H. Versnel and N. A. Kowalski, Organization of Primary Auditory Cortex evident in responses to rippled complex sound stimuli, *Neurosc. Meeting, November*, *19*, 581.9, 1993.

Shannon, R., Personal correspondance, 1992.

Slaney, M. and R. F. Lyon, Perceptual Pitch Detector, *Proc. ICASSP*, *1*, 357–360, 1990.

Vranić-Sowers, S. and S. A. Shamma, Representation of spectral profiles in the auditory system. II: Detection of spectral peak shape changes, *J. Acoust. Soc. Am. (submitted)*, 19xx .

Wang, K. and S. A. Shamma, Self-normalization and noise-robustness in early auditory representations, *IEEE Trans. Speech and Audio Processing*, *(to appear) July*, 1994.

Yang, X., K. Wang and S. A. Shamma, Auditory representations of acoustic signals, *IEEE Trans. on Information Theory*, *38*, 824–839, 1992.

Yost, W. A. and R. Hill, Models of the pitch and pitch strength of ripple noise, *J. Acoust. Soc. Am.*, *66(2)*, 400–410, 1979.